Empowering Adaptive Digital Game-Based Language Learning for Under-Resourced Languages Through Text Analysis

Elaine Uí Dhonnchadha[1], Sally Bruen[1], Liang Xu[2], Monica Ward[2]
[1] Trinity College Dublin, [2] Dublin City University
[1] Centre for Language and Communication Studies, Trinity College Dublin, Dublin, Ireland
[2] School of Computing, Dublin City University, Dublin, Ireland
{uidhonne, sbruen}@tcd.ie, liang.xu6@mail.dcu.ie, monica.ward@dcu.ie

Abstract

This study explores Cipher, an adaptive language learning game tailored for the under-resourced Irish language, aimed mainly at primary school students. By integrating text analysis techniques, Cipher dynamically adjusts its difficulty based on the player's language proficiency, offering a customised learning experience. The game's narrative involves decoding spells to access Irish myths and stories, combining language learning with cultural elements. Development involved collaboration with educators to align the game content with curriculum standards and incorporate culturally relevant materials. This paper outlines the game's development process, emphasising the use of text analysis for difficulty adjustment and the importance of engaging, educational gameplay. Preliminary results indicate that adaptive games like Cipher can enhance language learning by providing immersive, personalised experiences that maintain player motivation and engagement.

Keywords: adaptivity, text analysis, digital game-based language learning

1. Introduction

This paper focuses on the use of Natural Language Processing (NLP) to enhance player engagement in the *Cipher* language learning game. We do this by adapting the level of challenge in the game in response to the player's progress and proficiency. The scenario in the Cipher game is that an evil creature is putting stories and myths under a spell and making them unreadable, and the player's mission is to discover what spells were used and thwart the evil creature. In order to present stories and myths of an appropriate level we need to be able to rank and grade the stories. As this language learning game is being developed for an under-resourced language, in this case Irish, there are no tools for automatically ranking and grading Irish language texts, to our knowledge.

The layout of the paper is as follows. Section 2 gives background to learning Irish in Ireland, and introduces the benefits of educational games for language learning and the factors that influence their effectiveness. Section 3 gives an overview of *Cipher – Faoi Gheasa,* an adaptive educational game. Section 4 highlights our co-creation methodology for game development. Section 5 focuses on the development of text analysis tools for Irish which allow us to rank the game texts and thereby ensure that the level of challenge is adaptive to the individual player's proficiency.

2. Background

2.1 Irish language: background and challenges

The key features of successful language learning are engagement, motivation and early success. These factors are particularly important for first-time language learners who do not know what to expect when they start their language learning journey. This is the case for learners of Irish at primary school level in Ireland. Those who have prior experience of learning a language know that it can be challenging at times and requires a degree of resilience and that it is important to persevere when difficulties are encountered. For most L1 (first language) English primary school students in Ireland, their first exposure to another language is when they start learning Irish in primary school. This critical first experience can have an impact on future language learning. For example, if a student has a negative experience learning Irish, this could impact on their enthusiasm or self-confidence in learning another language. For this reason, it is very important that students have a positive experience of learning Irish. However, this is not always the case. In the majority of schools, teachers are not L1 speakers and lack confidence in their Irish ability. There are limited digital resources available for teaching and learning the language. Also, Irish plays a complex socio-cultural role in Ireland, where there is a tension between cultural value and economic value.

2.2 Game-Based Language Learning

Adaptive educational games blend the engagement of gaming with personalised learning, creating environments that motivate and immerse learners. Research (Peirce et al, 2008) argues for the intrinsic motivation these games offer, emphasising their capacity to adapt without sacrificing engagement. The innovation in non-invasively adapting games shows a balance between providing educational benefit and maintaining a game's immersive appeal.

According to Acquah and Katz's (2020) systematic review of empirical evidence for the effectiveness of digital games on second language learning, the main game features that influence outcomes are ease-of-use, challenge (i.e. being in the zone of proximal development), rewards and feedback, control or autonomy, goal-orientation, and interactivity. In this

Games and NLP @LREC-COLING-2024, pages 6–13

paper we focus on the first two features: ease-of-use and zone of proximal development (ZPD) (Vygotsky, 1978) as a means of making the game more engaging and adaptive to the users.

In striving to achieve the ZPD, since learning happens when the game is challenging but not too challenging, it is important that adapting educational content happens in a way that does not disrupt the gaming experience. This involves innovative approaches to personalising learning challenges and providing meta-cognitive support, ensuring that adaptations are non-invasive and augmentative rather than interruptive. Kickmeier-Rust et al. (2008) further highlight the role of micro-adaptivity in assessing competencies within games seamlessly, contributing to an improved learning and gaming experience compared to non-adaptive counterparts.

Adaptive educational games offer a promising avenue for enhancing learning through personalised, engaging experiences. The careful integration of adaptivity, aimed at bolstering motivation without compromising immersion, is pivotal. This approach not only enriches the educational potential of games but also preserves their intrinsic appeal as engaging and immersive entertainment mediums.

In Cipher, adaptivity is implemented by varying the number of Ciphers (spells) on each page, as well as adjusting the page's text length and the difficulty level of the text.

3. Cipher – an adaptive game with a purpose

Cipher, an adaptive language learning game, epitomizes the fusion of digital innovation and pedagogical strategy to enhance engagement and efficacy in learning under-resourced languages, notably Irish (Xu, 2023). The game focuses on engaging learners with tasks that promote incidental language learning through gameplay, centred around correcting language errors in the guise of magic spells. This approach not only makes learning interactive but also culturally resonant by embedding Irish mythology into the gameplay.

Acquah and Katz (2020) remarked on the lack of focus on cultural competence in the studies they reviewed, and recommended more research into understanding how digital learning games can be used to improve sociocultural outcomes. Attention to the 'spirit of the language' is a core feature of the Cipher language learning game, with a focus on folk tales and mythology which are culturally specific.

Content creation is carried out in collaboration with teachers ensuring that the game's educational material is aligned with curriculum standards, further enhancing its pedagogical value.

Figure 1: A screenshot of Cipher

Cipher's adaptability, which is also underscored by its modular design, showcases its potential for broader application across different linguistic contexts. Its development journey illustrates the power of combining engaging game mechanics with robust educational content, making it a model for future digital language learning tools.

4. Co-creation and Ease-of-use

The development of Cipher underscores the importance of a co-creation approach to educational game development, as it harnesses the collective expertise of a multidisciplinary team in crafting a digital game-based language learning resource. This collaboration is vital, as aligning the game with educational curricula is a critical factor for its acceptance and use in schools as highlighted by Ward (2007).

Given the challenges of assembling a comprehensive team for under-resourced languages, Cipher's development adopted a pragmatic strategy, utilising a core team consisting of a game designer, an NLP researcher specialising in Irish, a Computer Assisted Language Learning (CALL) researcher, and a primary school teacher. This team effort was essential in ensuring the game's relevance and suitability to learners' needs by incorporating curriculum-aligned content and simplifying language ensuring enhanced accessibility (Ward, 2023). An important advantage of this co-creation and curriculum-aligned approach is that the game is more suited to the needs of learners highlighting the necessity of diverse expertise in developing such resources.

Some enhancements to the Cipher game include the selection of texts at the correct level for the learners (see section 5) and the use of simpler and shorter sentences in the texts. While these may seem like obvious elements, it was essential to have confirmation from the teacher that these were necessary changes to the game. This co-creation approach, combined with a focus on ease-of-use and relevance for students enabled the development of an enhanced version of Cipher that would otherwise not have been possible.

5. Text analysis for adaptive learning

Xia et al (2016) note that while most studies of text readability are carried out by native speakers, L2

learners can perceive text comprehensibility very differently. Therefore a system such as Cipher that adapts to the L2 learner's perception of text difficulty based on their performance in the game will better facilitate language learning.

In order to provide texts of an appropriate level for adaptive learning it is necessary to be able to classify texts according to their relative complexity. In relation to second language acquisition, Michel (2017) notes that language complexity can be defined in a number of ways. It can be thought of as a) developmental complexity, i.e. the order in which linguistic structures are typically acquired, b) cognitive complexity, the subjective difficulty of a language feature from the perspective of the learner and c) linguistic complexity, i.e. objective complexity of forms, meanings and form-meaning mappings (Michel, 2017). In this section we focus on linguistic complexity and the features that can be used to determine the complexity of a text. In the literature, a wide variety of measures have been used to calculate linguistic complexity using a combination of lexical, grammatical and discourse features.

Regarding lexical measures, Lu (2012) describes lexical richness as a combination of lexical density (ratio of content words to total words), lexical variation or diversity (ratio of different words to total words) and lexical sophistication (proportion of advanced or sophisticated words in a text, using frequency lists or specialised word lists). Lexical familiarity/unfamiliarity (Collins-Thompson, 2014) is a similar concept to lexical sophistication.

Grammatical complexity can be characterised in terms of syntactic variation and sophistication, and most measures are based on mean length of sentence/utterance or on structural analysis (Lu, 2010; Vajjala, and Meurers, 2012).

Discourse measures of text quality include text cohesion, discourse relations and entity chaining (Pitler & Nenkova, 2008; Feng et al, 2010). Measures of discourse complexity that are associated with comprehension difficulty include propositional idea density (Kintsch & Keenan,1973; Brown et al, 2008). A text with greater propositional density is more difficult to comprehend than a text with lower propositional density. In this measure, predicates (e.g. verbs) modifiers (adjectives, adverbs, qualifiers) and conjunctions are considered to be indicative of propositions.

5.1 Text Analysis for Irish

Much of the research to date on text analysis and complexity measures has focussed on languages such as English (Vajjala & Meurers, 2012) while under-resourced languages such as Irish have not been researched to the same extent. In this section we calculate lexical, grammatical and discourse measures and investigate their usefulness as predictors for ranking and grading Irish texts. Given the small amount of pre-graded material available for an under-resourced language such as Irish, we choose to do a detailed linguistic analysis of the available material. This will satisfy our immediate need for grading texts for Cipher as well and help to generate graded material for machine learning methods.

We calculate lexical density as the ratio of content words to total words, lexical diversity includes corrected type-token ratio, and for grammatical complexity we use average sentence length, following findings for Irish in (Ó Meachair, 2019; Uí Dhonnchadha et al, 2022).

As noted in (Pitler and Nenkova, 2008) the vocabulary used in a text largely determines its readability, therefore we determine lexical sophistication (or familiarity) as the percentage of words that are among the N most frequent in a range of frequency bands using frequency word lists.

As regards discourse, since we are interested in reading comprehension difficulty rather than assessing text quality we investigate propositional idea density as described in (Brown et al, 2008).

5.2 Data and Methodology

We carry out a detailed analysis of pre-graded reading materials to develop predictive measures for ungraded material, i.e. to grade new stories for use in the Cipher game.

5.2.1 Data selection and preparation

Two sets of graded materials are used: *Taisce Tuisceana*[1] (TT), a collection of Irish reading comprehension texts for primary school students and *Séideán Sí*[2] (SS) a series of textbooks for *Gaeltacht* and Gaelscoileanna – Irish-medium primary schools in Ireland.

For initial exploration, we use ten samples of fiction from Taisce Tuisceana (TT). These texts are graded3 from A to E with A-C for Key Stage 1 (1st & 2nd class) and D-E for Key Stage 2 (3rd and 4th class). For increased robustness, we asked 5 Irish speakers to rank the texts in order of increasing difficulty based on their intuition, and we took the average of these ratings. While there was some variation within categories, there was very little variation between the categories, as shown in Table 1,which provided added confidence in the ranking.

[1] https://ccea.org.uk/learning-resources/taisce-tuisceana

[2] https://seideansi.ie/

[3] Graded A for 1st & 2nd class, B&C for 3rd and 4th class and D&E for 5th and 6th class on www.tairseachcogg.ie/

Taisce Tuisceana	Cat.	Class	R1	R2	R3	R4	R5	Ave Rank	StDev
A06	A	1st, 2nd	1	1	1	1	1	**1.00**	0.00
B22	B	1st, 2nd	3	2	2	2	2	**2.20**	0.45
B20	B	1st, 2nd	2	5	4	6	3	**4.00**	**1.58**
C23	C	1st, 2nd	5	3	3	3	6	**4.00**	**1.41**
C08	C	1st, 2nd	4	4	5	5	5	**4.60**	0.55
D05	D	3rd, 4th	6	7	6	4	4	**5.40**	**1.34**
D08	D	3rd, 4th	8	6	7	7	7	**7.00**	0.71
D21	D	3rd, 4th	7	8	9	9	8	**8.20**	0.84
D22	D	3rd, 4th	9	9	8	10	9	**9.00**	0.71
E02	E	3rd, 4th	10	10	10	8	10	**9.60**	0.89

Table 1: *Taise Tuisceana* texts and grading

The second dataset *Séideán Sí* (SS), is larger and more comprehensive. It consists of textbooks which are pre-graded for primary school students from first class (ages 6-7) to sixth class (ages 11-12). From this collection we selected 86 samples of fiction, 10% of which (8 files) were set aside for testing and the rest were used to develop grading measures.

Stage	Class	Age	Files	Words
Keystage 1 (6-8 years)	1st Class	7-8	30	5721
	2nd Class	8-9	24	15679
Keystage 2 (8-11 years)	3rd Class	9-10	6	15539
	4th Class	10-11	10	23585
	5th Class	11-12	9	28894
	6th Class	12-13	7	40793
TOTAL			86	130211

Table 2: *Séideán Sí* texts

The PDF files which were converted to text and cleaned, were then part-of-speech (POS) tagged and lemmatized using existing tools for Irish (Uí Dhonnchadha and Van Genabith, 2006). This rule-based POS tagger achieved an average accuracy of 97% on a random selection of eight *Séideán Sí* texts covering all levels. Along with these stories, a 10K frequency wordlist was extracted from the fiction part (6 million words approx.) of the New Corpus of Ireland[4] (NCI).

5.2.2 Calculation of Text Statistics

Text statistics were obtained for each story using Python programs which process each POS-tagged file to get the required lexical, grammatical and frequency measures.

For the lexical measures, TTR (type token ratio), WTR (word type ratio) and CTTR (corrected type token ratio) were calculated by using formulae in (Vajjala & Meurers, 2012; Lu, 2012). For the grammatical measure, WDSEN (average number of words per sentence) was calculated along with various other text statistics as shown in Table 3. For word familiarity (sophistication) measures, we compared the word types in each text with the frequency wordlist and calculated the number of words that were within frequency ranges 0-100, 100-300, 300-500, 500-1000, 1000-2000, 2000-3000, 3000-4000, 4000-5000, 10000+. Each story was labelled based on its year group e.g. first class stories were labelled '01_' followed by the story name while sixth class stories were labelled '06_' followed by the story name.

Finally, propositional idea density was calculated using POS tags for verbs, copula, adjectives, adverbs and conjunctions.

Feature Counter

The first tool calculates text statistics such as number of sentences, tokens, types, words, lemmas, longest sentence, and POS totals (using the first two characters of the PAROLE[5] POS tag e.g. Np, Vm, Pp, Nv, etc.) and nouns with genitive case. Table 3 gives an example of the output from this tool.

Lexical and Grammatical Calculator

The second tool calculates lexical and grammatical measures for the files. These measures were TTR, WTR, CTTR and WDSEN. Table 4 gives an example of the output from this tool.

FILENAME	TTR	WTR	CTTR	WDSEN
01_C1_An	0.533333	1.875	3.577709	6.923077

Table 4: Output from the Lexical and Grammatical calculator tool

Frequency Range Tool

The third tool calculates the frequency measures for each file by using a wordlist of the 10,000 most frequently used words in Irish fiction. For every word in each file, if the word was within the frequency list, the appropriate frequency range was incremented, otherwise the 10K+ range was incremented. Table 5 gives an example of the output from this tool.

5.3 Results

Graphs were used to visualise the results. CTTR and average sentence length prove to be the most predictive measures, showing increasing values for first to sixth class texts. Figure 2 shows the range of CTTR values for *Séideán Sí* (SS) files (labels 01 – 06), and *Taisce Tuisceana* (TT) files (labels A – E). Figure 3 shows the average sentence length for SS and TT files.

[4] https://corpas.focloir.ie/

[5] https://www.scss.tcd.ie/~uidhonne/parole.htm

FILENAME	sen_count	tokens	types	word_count	ave_sen_len	lemtypes	gen_count	Np	Pr	Av
01_C1_Am	13	109	48	90	6.92307692	40	3	2	6	0

Table 3: Output from the Feature counter tool

FILENAME	TYPES	100FREQ	300FREQ	500FREQ	1000FREQ	2000FREQ	3000FREQ	4000FREQ	5000FREQ	10000FREQ	10KplusFREQ
01_C1_An	49	24	9	2	2	3	2		1	3	3

Table 5: Output from Frequency tool

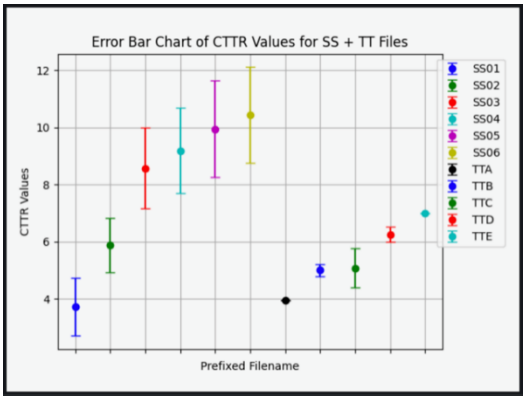

Figure 2: CTTR values for SS, TT texts

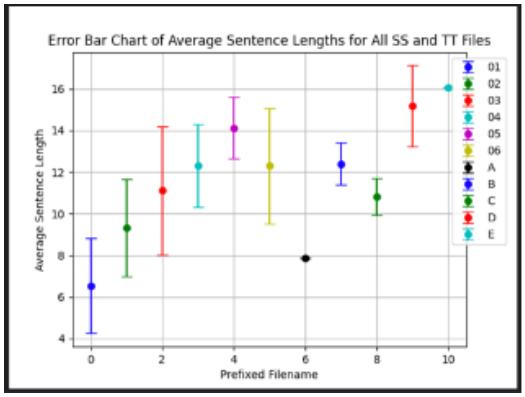

Figure 3. Average sentence length values for SS (01-06) and TT(A-E) files

Figures 4 and 5 show each SS and TT text groups and their percent of types in each frequency range from 0-100 to 10K+. In this Zipf-like curve the proportion of words in the 1-100 frequency range appears to be the most useful frequency range for predicting text level.

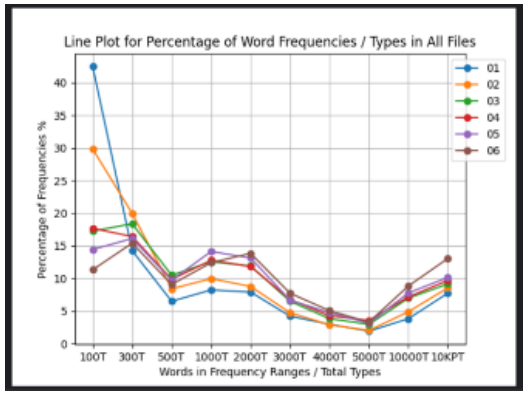

Figure 4: Type/Frequency ranges of SS texts

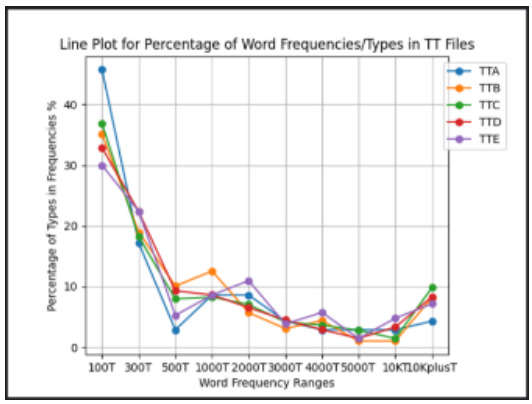

Figure 5: Type/Frequency ranges of TT texts

The graphs show that the measures of CTTR, average sentence length and frequency for the pre-graded texts are behaving as one would expect, with an increase from first class texts to sixth class texts.

In Figure 2 we see that the TT files fall within the range of SS 1st to 3rd class (01-03). In Figure 3 the average sentence length for TT files fall within the range of 1st to 6th class and higher. It is also apparent that higher level texts do not always have a greater average sentence length, as can be seen from 05 and 06 texts as well as B and C texts. This suggests that a composite measure is required as there is often a trade-off between lexical complexity and grammatical complexity (Graesser and McNamara, 2011).

By plotting the ungraded texts alongside these graded texts, we can infer a grading for the texts based on where they are in the plots in relation to the SS and TT files.

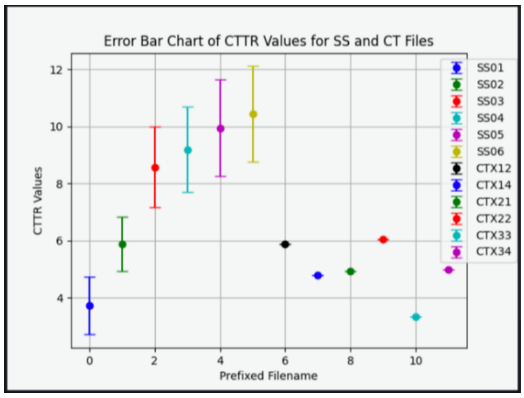

Figure 6: CTTR for SS and ungraded Cipher Texts (CT)

In Figure 6 we show CTTR for six ungraded Cipher Texts (CT) in comparison to the SS texts, and in Figure 7 we show average sentence length for the same six ungraded Cipher Texts (CT) in comparison to the SS texts.

The Cipher texts comprise of two Irish mythology texts (CTX12, CTX14), two Dúchas[6] Irish folklore texts (CTX21, CTX22) and two international fairy tales (CTX33, CTX34).

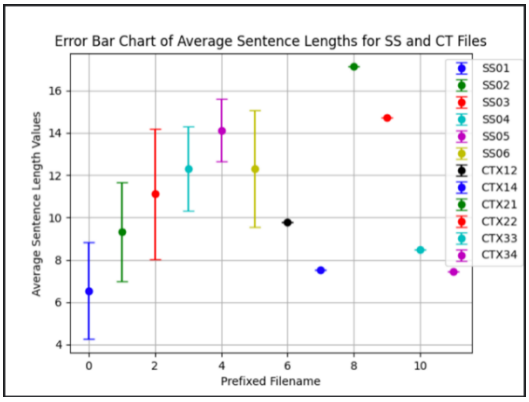

Figure 7. Average sentence length values for SS and ungraded Cipher Texts (CT)

As we can see from Figures 6 and 7 the mythology texts (CTX12, CTX14), fall within the 2nd and 3rd class range, the Dúchas texts (CTX21, CTX22) are lexically in the 2nd class range but are in the 5th and 6th range and above in terms of sentence length, while the fairy tales are in the 1st and 2nd class ranges. These results are as expected. In particular the results for the Dúchas texts are interesting, given that these texts were written by L1 Irish children in 5th 6th class in the 1930's in a National Folklore Collection project.

5.3.1 Composite predictive measures

These visual results are useful, but we need to able to automatically assign a grade in the range 1-6 to ungraded texts. Predicting the grade level of ungraded texts, was carried out in an incremental manner. We began with one feature, CTTR. From our pre-graded SS texts (Fig. 2), we could estimate value ranges for this measure for each age group. For example, we could see the majority of the first class files had a CTTR value between 1 and 4, so we set our first range to be 1 to 4, while range 2 which correlated with second class was 4 to 7.6, range 3 for third class was 7.6 to 8.6, range 4 for fourth class was 8.6 to 9.5, range 5 for fifth class was 9.5 to 13.5 and range 6 for sixth class was 13.5 plus. With every file, the CTTR value was compared with these ranges and when a value landed in a particular range, that range number (1 to 6) was returned to give the CTTRRangeValue. From using this measure alone, we could see there were multiple outliers in each age group that were giving a higher range value than expected so to help account for those outliers, the WDSEN measures was added.

The same principles were followed for this measure by having different ranges for this measure. With these two measures, an average was taken from the CTTRRangeValue and WDSENRangeValue to give a rounded predicted grade level. The predicted grade level was still not overly accurate for all six grade levels. Therefore the measure '100T' was added to the calculation following the same format and having its own range values. This measure calculates the percentage of types in a file that are within the first one hundred words of our frequent word list. The TypeRangeValue was added with the two other range values to get an overall average which still was unsatisfactory, leading to the addition of the lemma type count to the calculation. A set of ranges were set for lemma types and the value was added to the previous three features.

Feature	Exact match	Within 1 grade level
CTTR	58.76%	31.2%
+WDSEN	55.29%	36.4%
+100T'	59.4%	30.58%
+Lemma types	60.0%	34.12%

Table 6 Initial grade prediction results

The average from this had a more satisfactory result with 60% of the predictions being exactly right with grade level, while 34.1% of the predictions being out by one grade level. This is a good result with over 90% predictions being within one of the correct grade. See Table 6 for details.

Held out SS Text	Predicted Grade	Actual Grade
IgCéinisigCóngar.txt	5	6
AnMúinteoirNua.txt	4	3
Céhiadseo.txt	1	1
AnNathairagusnaSpléaclaí.txt	3	3
SinScéalEile.txt	5	5
MurachanTraenáil.txt	2	2
SciobAgusAnChuileog.txt	1	1
AChaitlínMí-abha.txt	3	3

Table 7 grade predictions for held out SS texts

To test this formula on unseen texts, a random selection of 10% held-out SS files were used to check accuracy levels, and for these 8 files, 70% of the predictions were exactly right with 30% of the predictions being out by one grade level. See Table 7 for details. The predictive grading measure is used to assign a grade to the ungraded Cipher texts. These results are in keeping with the visual results in Figures

[6] National Folklore Collection - Dúchas https://www.duchas.ie/en/cbes/stories

5 and 6 and with the team members expectations. The Dúchas texts (CTX21, CTX22) are predicted to be suitable for 3rd and 4th grade which reflects a combination of 2nd grade lexical features and 5th and 6th grade sentence length.

These automatic grading results are promising and we intend to verify them by carrying out more extensive testing on unseen, manually graded texts.

A limitation of this study is the small sample size. Also we limited our study to only fiction texts, therefore it is unclear how well the predictions would generalise to other types of text.

6. Conclusion

The development and implementation of Cipher highlights an advancement in digital game-based language education, particularly for under-resourced languages. Through analysis techniques and a focus on adaptivity, Cipher offers a personalised learning experience that dynamically adjusts to the learner's proficiency level, ensuring that the challenge remains within the learner's zone of proximal development. The integration of Irish culture and mythology not only enhances engagement and motivation but also fosters a deeper connection with the language. Preliminary findings suggest that such adaptive educational games hold promise in transforming language learning by making it more engaging, effective, and enjoyable. By addressing the unique challenges associated with teaching and learning less commonly taught languages, Cipher paves the way for future research and development in the field of digital game-based language learning, and in this case has inspired the development of new NLP text analysis tools for Irish.

In a collaborative co-creation model we aim to increase player engagement and educational effectiveness in Cipher by focusing on enhanced adaptivity in response to player progress in the game. We also aim to increase engagement by improving ease-of-use through more gradual transitions between levels and the use of simpler and shorter sentences in the texts in the earliest levels.

This study exemplifies the potential of combining gaming technology and CALL research with linguistic analysis and cultural content to enrich language education, offering valuable insights for educators, developers, and researchers who aim to enhance language learning outcomes through innovative digital solutions.

7. Bibliographical References

Acquah, E.O. and Katz, H.T. (2020). Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review. Computers & Education, Volume 143.

Brown, C., Snodgrass, T., Kemper, S.J., Herman, R. & Covington, M.A. (2008). Automatic measurement of propositional idea density from part-of-speech tagging. Behavior Research Methods 40, 540–545.

Collins-Thompson, K. (2014). Computational assessment of text readability: A survey of current and future research. ITL - International Journal of Applied Linguistics, 165(2):97–135.

Feng, L., Jansche, M., Huenerfauth, M., & Elhadad, N. (2010). A comparison of features for automatic readability assessment. COLING 2010 Poster Volume, pp. 276-284.

Kickmeier-Rust, M. D., Hockemeyer, C., Albert, D., & Augustin, T. (2008). Micro adaptive, non-invasive knowledge assessment in educational games. In 2008 Second IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning (pp. 135-137).

Kintsch, W.& Keenan, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology,5:3, pp 257-274.

Lu, X. 2010. Automatic analysis of syntactic complexity in second language writing. International Journal of Corpus Linguistics, 15(4):474–496.

Lu, X. (2012), The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. The Modern Language Journal, 96: 190-208.

Michel, M. (2017). Complexity, Accuracy and Fluency in L2 production. In S. Loewen, & M. Sato (Eds.), Routledge Handbook of Instructed Second Language Acquisition (pp. 50-68). (Routledge handbooks in applied linguistics). Routledge.

Ó Meachair, M. J. (2019). The Creation and Complexity Analysis of a Corpus of Educational Materials in Irish (EduGA). PhD, University of Dublin, Trinity College.

Peirce, N., Conlan, O. and Wade, V. (2008). Adaptive educational games: Providing non-invasive personalised learning experiences. In 2008 second IEEE international conference on digital game and intelligent toy enhanced learning, pp. 28-35.

Pitler, E. and Nenkova, A. (2008). Revisiting readability: A unified framework for predicting text quality. In Proceedings of the 2008 conference on empirical methods in natural language processing, pages 186–195.

Proceedings of the 2008 conference on empirical methods in natural language processing, pages 186–195.

Uí Dhonnchadha, E., Ward, M. & Xu, L. (2022). Cipher – Faoi Gheasa: A Game-with-a-Purpose for Irish. In Proceedings of the CLTW 4 @ LREC2022 , Marseille, June 2022, pp 77–84

Uí Dhonnchadha, E. & Van Genabith, J. (2006). A Part-of-speech tagger for Irish using Finite-State Morphology and Constraint Grammar Disambiguation. LREC 2006, May 2006 Genoa.

Vajjala. S. (2022). Trends, Limitations and Open Challenges in Automatic Readability Assessment Research. In Proceedings of the Thirteenth

Language Resources and Evaluation Conference, pages 5366–5377, Marseille, France.
Vajjala, S. & Meurers, D. (2012). On Improving the Accuracy of Readability Classification using Insights from Second Language Acquisition. 7th Workshop on the Innovative Use of NLP for Building Educational Applications, 2012 Monréal, Canada. ACL, 163-173.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.
Ward, M. (2007). The integration of CL resources in CALL for Irish in the primary school context (Doctoral dissertation, Dublin City University).
Ward, M., Uí Dhonnchadha, E., McGarry J. and Xu, L. (2023). Co-creating CALL content - does it work?Goldilocks compromise or Cruella chaos? In CALL for all Languages - EUROCALL 2023 Short Papers. 15-18 August 2023,University of Iceland, Reykjavik.https://doi.org/10.4995/EuroCALL2023.2023.16961
Xia, M., Kochmar, E., & Briscoe, T. (2019). Text readability assessment for second language learners. arXiv preprint arXiv:1906.07580
Xu, L., Uí Dhonnchadha, E and Ward, M., Harnessing the Power of Images in CALL: AI Image Generation for Context-Specific Visual Aids in Less Commonly Taught Languages, EUROCALL2023, Reykjavík, Iceland, 15-18 August, 2023